170-WP-007-004

CEOS Interoperable Catalogue System (ICS) System Design Document (SDD)

ECS White Paper

March 1997

Prepared Under Contract NAS5-60000

RESPONSIBLE ENGINEER

George Percivall /s/ 3/14/97

George Percivall Date
EOSDIS Core System Project

SUBMITTED BY

Paul Fingerman /s/ 3/14/97

Paul Fingerman Date
EOSDIS Core System Project

Hughes Information Technology Systems
Upper Marlboro, Maryland

Committee on Earth Observation Satellites
Working Group on Information Systems and Services

Interoperable Catalogue System (ICS)
System Design Document (SDD)

CEOS

Working Group on Information Systems and Services

Protocol Task Team

Doc. Ref.: CEOS/WGISS/PTT/SDD

Date: March 1997

Issue: Version 1.2

AUTHORITY

Issue:	Version 1.2
Date:	March 1997
Location:	PTT Teleconference

This document has been approved for publication by the Protocol Task Team of the Committee on Earth Observation Satellites (CEOS) and reflects the consensus of the Protocol Task Team technical panel experts from the CEOS member agencies.

This document is published and maintained by:

Gerhard Triebnig (CEOS PTT Chair)
European Space Agency/ESRIN
Via Galileo Galilei
00044 Frascati
Italy

E-mail: *gerhard.triebnig@esrin.esa.it*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DOCUMENT STATUS SHEET

Version	Date	Comments
0.1	May 1996	Interoperable Catalogue Systems and Services System Architecture Document (ECS White Paper 170-WP-007-001). Annotated Outline of SAD prepared for PTT Meeting, May 1996, Tokyo.
1.0	September 1996	Document renamed ICS SDD. RIDs from Tokyo PTT meeting incorporated. Document prepared for PTT Meeting, September 1996, Annapolis. Final scope and purpose. Initial technical completeness review.
1.1	November 1996	RIDs from Annapolis PTT meeting incorporated. Document prepared for PTT Meeting, December 1996, Ispra. Document to be reviewed by PTT using formal RID process.
1.2	March 1997	RIDs from Ispra PTT-4 meeting incorporated. Document prepared for ratification by PTT.

DOCUMENT CHANGE RECORD

Change No.	Date	Changes in Version 1.2 from Version 1.1
1	January 1997	Incorporated RIDs from PTT-4 meeting in Ispra.
2	January 1997	Consistency with CIP Specification on compliance statement (Section 1.1), attribution (Section 1.2), and catalogue interoperability overview (Section 2).

This page intentionally left blank.

1. INTRODUCTION

1.1 SDD Purpose and Scope

The CEOS Protocol Task Team (PTT) is developing a suite of documents as defined in the PTT Development Plan [R1][*]. The ICS User Requirements Document (URD) [R2] is where the CEOS agencies bring in their needs for functionality, constraints, etc. The PTT accepts input as requirements added to the URD. It is the PTT activity to interpret those requirements into the CIP Specification [R3] and the ICS SDD. As a result, if groups comply with the CIP Specification and the SDD, then they will be compliant with the URD. The SDD may generate new requirements, derived from doing the system design, which are made explicit on occasion. The SDD is written for a federation of CEOS agency systems, however other data provider federations could use the SDD as a template design for their system.

The ICS SDD defines the elements and interfaces which comprise the CEOS ICS. This document also defines relationships to the CEOS Network which is under development by the CEOS Network Subgroup. Assumptions are made in this document about existing systems which will interface to ICS, e.g., existing agency systems will persistently store user orders. These assumptions cannot be required of the agency systems but represent the SDD's approach to the interface with ICS. The last section of the SDD defines a range of site configurations relative to the SDD contents. The ultimate purpose of the SDD is to define a system which can be implemented and operated by the CEOS federation to provide data and services to users.

The purpose of ICS is provided in Section 2. This section defined the purpose of the SDD.

1.2 Organization of the SDD

The SDD is organized into various views of the ICS. This approach is based on guidelines for SDDs, e.g., [R6] and existing SDDs, e.g., [R7]. After a discussion of the PTT approach to catalogue interoperability (Section 2), the SDD provides the following views of the ICS:

- Functional Framework (Section 3) - overall application view of the computing elements which comprise the functional architecture of ICS. The ICS Framework provides several configurations for ICS implementation at sites.
- Data View (Section 4) - provides an overall data framework indicating what data is held by the various ICS elements. Several data components are defined in the Retrieval Manager in order that ICS operations are performed correctly across ICS sites.
- Communications View (Section 5) - describes two models of the ICS design related to communications and networks. First a communication protocol stack is defined placing CIP on a TCP/IP stack. Next, ICS reliance on CEOSnet is discussed including identifying sites at which ICS Retrieval Managers are initially intended to be located and the traffic between the sites.

[*] Citations for references are provided in Section 1.5.

- Security (Section 6) - This section provides a discussion of the security design for ICS. First, a security assessment of ICS is provided in terms of security assets, vulnerabilities and threats. Based on the assessment, the ICS security control methods are described. The methods are grouped into three categories: administration, physical and computing controls.
- Systems Management (Section 7) - This section provides a list of topics which can serve as a basis for beginning the System Management design in the next ICS release.
- Architecture Verification (Section 8) - This section provides several methods for evaluating the system design. Through a series of scenarios, a definitive identification of inter-element interfaces, and distributed query performance estimates, this section demonstrates the completeness of the system design presented in earlier sections.
- ICS Compatibility and Configurations (Section 9) - This section identifies several canonical site configurations and what is required at the sites to be considered ICS compatible.

1.3 ICS/PTT Development Process

This document was developed by the CEOS Protocol Task Team (PTT). The PTT is part of the WGISS Access Subgroup within CEOS. The lead agency for compilation of PTT inputs and for final preparation of this version of the document was NASA. A complete list of organizations participating in the PTT is provided in the PTT Terms of Reference (http://ceos.ccrs.nrcan.gc.ca/taskteam/cip.html)

1.4 Glossary

1.4.1 Acronyms

The following acronyms are used in this document:

ADD	Architecture Design Document
ANSI	American National Standards Institute
APDU	Application Protocol Data Unit
AS	Access Subgroup (Part of CEOS-WGISS)
ASN.1	Abstract Syntax Notation.1
BER	Basic Encoding Rules
BNSC	British National Space Centre
CA	Certification Authority
CCRS	Canada Centre for Remote Sensing
CCSDS	Consultative Committee for Space Data Systems
CDB	Collection Data Base
CEO	Centre for Earth Observation (European Commission)
CEO-ES	CEO- Enabling Services
CEOS	Committee on Earth Observation Satellites
CINTEX	Catalogue INTeroperability EXperiment

CINTOPS	Catalogue INTeroperability OPerationS
CIP	Catalogue Interoperability Protocol
CM	Configuration Management
CMT	Collection Management Tools
CNES	Centre National d'Etudes Spatiales (France)
CSIRO	Commonwealth Scientific and Industrial Research Organisation (Australia)
DB	Data Base
DBMS	Data Base Management System
DEM	Digital Elevation Map
DIF	Directory Interchange Format
DLR	Deutsche Forschungsanstalt fur Luft-und Raumfahrt
ECS	EOSDIS Core System
EO	Earth Observation
EOC	Earth Observation Center (NASDA)
EOS	Earth Observing System
EOSDIS	Earth Observing System Data and Information System (NASA)
ESA	European Space Agency
ESRIN	European Space Research Institute (ESA)
GCMD	Global Change Master Directory
GSFC	Goddard Space Flight Center (NASA)
HITS	Hughes Information Technology Systems, Inc.
HTML	Hyper Text Mark-up Language
HTTP	hypertext transfer protocol
ICS	Interoperable Catalogue System
ID	IDentifier
IP	Internet Protocol
IRE-RAS	Institute of Radio Engineering. - Russian Academy of Science
ISO	International Standards Organization
LaRC	Langley Research Center (NASA)
LDAP	Lightweight Directory Access Protocol
MAC	Message Authentication Code
MACAO	Member Agency Control Authority Office
MCT	Monitoring and Control Tools
NASA	National Aeronautics and Space Administration (US)
NASDA	National Space Development Agency (Japan)
NOAA	National Oceanic and Atmospheric Administration (US)
NRSC	National Remote Sensing Center

NSRS	Natural Environment Research Council (BNSC)
OHS	Order Handling System
OSI	Open Systems Interconnection
PTT	Protocol Task Team (Part of CEOS- WGISS-AS)
PKI	Public Key Infrastructure
QA	Quality Assurance
RFC	Request For Comment
RM	Retrieval Manager
RMA	Retrieval Manager Administrator
RPN	Reverse Polish Notation
SATAN	Security Administrator Tool for Analyzing Networks
SDD	System Design Document
SE/RC	System Evolution/Release C (ECS)
SFDU	Standard Formatted Data Unit
SNMP	Simple Network Management Protocol
SST	Sea Surface Temperature
TBD	To Be Determined
TBR	To Be Resolved
TBS	To Be Supplied
TCP	Transmission Control Protocol
TN	Technical Note
UDB	User Data Base
UPS	User Profile System
UR	User Requirement
URD	User Requirements Document
URL	Uniform Resource Locator
URN	Uniform Resource Name
US	United States
USGS	U.S. Geological Survey
WGISS	Working Group on Information Systems and Services (Part of CEOS)
WWW	World Wide Web

1.4.2 Definitions

This section provides definitions of the terms related to ICS:

Archive	See definition in ICS URD [R2]
Catalogue System	See definition in ICS URD [R2]
Catalogue Translator	One of three types of ICS *Translators. Catalogue Translator* converts CIP messages into a data providers protocol for the services of Inventory, Directory, Browse, Guide and Ordering.
CEOS Agency Systems	The data provider systems of CEOS member agencies and their affiliates.
CIP Client	A software element composed of a *Presentation Layer*, a *Local User Management Layer*, and an *Application Layer.* Only the *CIP Client Application Layer* is part of ICS.
CIP Client Application Layer	Part of the *CIP Client* which deals directly with CIP including creating CIP messages and includes off-the-shelf Z39.50 communication software.
CIP Client Local User Management Layer	Part of the *CIP Client* which provides functionally for local data management, e.g. saving a result set, or converting result sets into orders.
CIP Client Presentation Layer	Part of the *CIP Client* dealing with how information is presented to the human user, including all issues related to HMI, as well as dealing with certain format specific issues, e.g., displaying browse imagery.
CIP Operation	Based on Z39.50 definition, an initiating CIP request message and the corresponding terminating response, along with intervening related messages. Multiple operations may occur within a *CIP Session*.
CIP Session	A set of CIP messages exchanged between an *Origin* and *Target* beginning with an initialization message and ending with a close message between which the *Origin* and *Target* maintain state information concerning the interaction. (A *CIP Session* is a Z39.50 Z-Association.)
CIP Message	A unit of information transferred between an origin and a target whose format is specified as a Z39.50 Application Protocol Data Unit (APDU) possibly containing CIP specific APDUs in the external portion of a Z39.50 APDU.
Collection	A grouping of item descriptors that have commonality. A collection consists of a number of attributes that describe the collective contents of the collection, the values of these attributes can then be searched on to select items of interest to the user. Collections also have members; these are the unique identifiers of the items that are grouped by the collection rather than their collective descriptions. As collection members can be identifiers of other collections, a hierarchy of collections and product/guide members can be established, therefore permitting a flexible and powerful organization of data.

Collection Management Tool	Used by the RMA for tasks involved with populating and maintaining the data in the Retrieval Manager. These tasks involve translating collection or directory information into CIP collection format and checking for valid entries.
Existing Agency Client	Software elements which interact with a data providers *Catalogue* using the data providers protocol.
Guide data	Data that is available to the user to enhance understanding of the EO data, spacecraft, instrument, etc., and hence make a detailed analysis of whether the product data will be of value for a particular application. Guide data may also contain information necessary for processing the product data further, such as calibration coefficients.
ICS Gateway	Software element which provide ICS access to a data providers catalogue, e.g., searches by *Existing Agency Clients* to a *Catalogue* are sent to a *Retrieval Manager* using an *ICS Gateway.*
Item Descriptor	Used by CIP to represent items. The descriptor is a set of attributes. Item descriptors are of the following types: product, collection, guide, user.
Monitoring and Control Tools	Provides the machine-to-machine interface for integrating the operations of the *Retrieval Manager* with the operations of SSM.
OHS Translator	One of three types of ICS *Translators. OHS Translator* converts CIP messages into a data providers protocol for the purposes of order specification, quotation, and request.
Order Handling System	That part of a data provider which provides services associated with ordering products and guide.
Origin	Based on the Z39.50 definition for origin, where an *Origin* is that portion of a client or middleware which may initiate a *CIP session* with a *Target*. ICS elements capable of acting as *Origins* are *CIP Clients, Retrieval Managers, ICS Gateway HTTP/CIP Gateways*.
Other CIP Based Federations	Groups of organizations other than CEOS using CIP to provide catalogue interoperability to data providers. These providers are interoperable with CEOS using CIP. These other federations may or may not use the ICS system design as a basis of the federation design.
Other Z39.50 Based Services.	Catalogue interoperability services provided using Z39.50 Version 3 but not necessarily CIP.
Primary Order	A CIP order between a *CIP Client* and a *Retrieval Manager.* (See *Secondary Order*)
Product data	A unique aggregation of data generated from information held in, or to be held in an archive (for predicted products). It can be located and retrieved by a user via CIP, possibly following further processing, such as map projection, sub-setting, band selection, etc., after or during extraction of the raw data as stored in the archive.
Retrieval Manager	See definition in ICS URD [R2]

Retrieval Manager Administrator	The human operator that performs all tasks needed to establish and maintain a *Retrieval Manager*. In practice this is more than one person as the tasks are various types: scientist for collection definition, data base expert for maintaining CDB, system operator for diagnosing and correcting operational activities, etc. For convenience purposes all of these tasks are performed by the *RMA.*
Secondary Order	A CIP order is created by an *Retrieval Manager* in response to a *primary order* and may be either to another *Retrieval Manager* or to an *OHS Translator*
Site System Management	That part of a data provider which provides coordinated, on-line management of the distributed processing environment.
Target	Based on the Z39.50 definition for target, where a *Target* is that portion of middleware and servers which accept requests for *CIP sessions* from an *Origin*. ICS elements capable of acting as *Targets* are *Retrieval Managers* and *Translators.*
Task Package	The set of attributes that describe an activity which is started by an Extended Services Request. Based on Z39.50 definition for a Task Package.
Translators	Software element which converts CIP into the protocols used by a data provider. Three *Translators* are identified in ICS: *Catalogue Translator, OHS Translator, UPS Translator.*
UPS Translator	One of three types of ICS *Translators. UPS Translator* converts CIP messages into a data providers protocol for the user information, e.g., address, and for authentication.
User Profile System	That part of a data provider which provides services associated with user information and authentication.
User Session	See definition in ICS URD [R2]. A user session results in at least on *CIP session* with a *Retrieval Manager* and perhaps other *CIP sessions* between the *Retrieval Manager* and other ICS elements.
Users	See definition in ICS URD [R2]

1.5 References

[R1] *Protocol Task Team (PTT) Development Plan,* CEOS/WGISS/PTT/Plan, Issue 1.0, 19 July 1996, Committee on Earth Observation Satellites /Protocol Task Team

[R2] *Interoperable Catalogue System (ICS) User Requirements Document (URD),* CEOS/WGISS/PTT/ICS-URD, Issue 2.2, to be ratified March 1997, Committee on Earth Observation Satellites /Protocol Task Team

[R3] *Catalogue Interoperability Protocol (CIP) Specification - Release B*, CEOS/WGISS/PTT/CIP, Issue 2.2, to be ratified March 1997, Committee on Earth Observation Satellites /Protocol Task Team

[R4] (reserved)

[R5] *ICS Collection Manual,* CEOS/WGISS/PTT/CM, Issue 1.0, planned for first quarter of 1997, Committee on Earth Observation Satellites /Protocol Task Team

[R6] *Technical Architecture Framework for Information Management (TAFIM), Volume 3: Architecture Concepts and Design Guidance*, Version 2.0, 30 June 1994, US Defense Information Systems Agency Center for Architecture.

[R7] *Summary of the ECS System Design Specification*, ECS Technical Paper Number 19400436, June 1994, Hughes Information Technology Systems (HITS) ECS Project Office. (document is available at http://edhs1.gsfc.nasa.gov/Info/sdr/tp19400436cat.html)

[R8] *(reserved)*

[R9] *Information Retrieval (Z39.50): Application Service Definition and Protocol Specification*, ANSI/NISO Z39.50-1995, Official Text, July 1995, Z39.50 Maintenance Agency.

[R10] *DPRS Technical Note/002*, LUK.502.EC21317/TN002, Issue 0.1, 20 September 1995, Logica UK Ltd.

[R11] *ICS Collection Technical Note*, CEOS/WGISS/PTT/CTN, Version 1.0, 8 July 1996, Committee on Earth Observation Satellites, Protocol Task Team

[R12] *Internetworking with TCP/IP: Volume 1, Principles, Protocols and Architecture*, Douglas E. Comer, Prentice-Hall, 1995.

[R13] *Z39.50 Information Retrieval Protocol in the Internet Environment*, Clifford A. Lynch . (Document is available at http://lcweb.loc.gov/z3950/agency/)

[R14] *ISO 8824 - Information Processing Systems* - Open Systems Interconnection - Specifications for Abstract Syntax Notation One (ASN.1), 1987

[R15] *ISO 8825 Information Processing Systems* - Open Systems Interconnection- Specification of Basic Encoding Rules for Abstract Syntax Notation One (ASN.1) 1987

[R16] *RFC 1340 Assigned Numbers*. Reynolds, J.K.; Postel, J.B. July 1992

[R17] *CEO Enabling Services Usage Scenarios Document*, CEO/ES/WP 2200/SD-1/184, Document Issue 1, July 1996, Centre for Earth Observation (CEO).

[R18] (reserved)

[R19] *ICS Query Performance Study Report*, under development by R. Howard, Hughes Information Technology Systems, to be delivered by January 1997.

[R20] *ICS Network Bandwidth Study Report*, under development by R. Howard, Hughes Information Technology Systems, to be delivered by January 1997.

[R21] *Queuing Systems*, Volume 1: Theory, Leonard Kleinrock, John Wiley, 1975.

[R22] *ECS User Scenario Functional Analysis*, NASA ECS White Paper, Document 194-00548TPW, Hughes Information Technology Systems, October 1994

[R23] T*he Directory - ITU-T Authentication Framework*, ITU-T (formerly CCITT), Recommendation X.509, 1994 (Version 3 extensions)

2. ICS DESIGN APPROACH

This section provides an introduction to the PTT approach to catalogue interoperability. This approach is consistent between the ICS SDD and the CIP Specification. This section also introduces the concept of Architecture Views which is the organizational approach to the presenting the ICS design in the SDD.

2.1 Introduction to Catalogue Interoperability

This section provides an overview of the Protocol Task Teams efforts in support of catalogue interoperability. The following topics are addressed in this section:

- CEOS and the Protocol Task Team
- Purpose and scope of catalogue interoperability
- Design approach: CIP space and ICS
- Collections data model
- CIP as a Z39.50 profile
- Browse data in CIP
- Product ordering and security
- Implementation of CIP and ICS

2.1.1 CEOS and the Protocol Task Team

The Committee on Earth Observation Satellites (CEOS) is comprised of international space agencies. CEOS promotes the interoperability of space agency catalogues through the definition and development of interoperability concepts.

As part of CEOS, the Protocol Task Team (PTT) collects requirements for an information retrieval and order protocol for EO data catalogue and access systems. The PTT establishes agreement on these requirements and produces detailed specifications on which system implementations can be based. These requirements and specifications provide a basis for interoperability using modern standard solutions and which respects CEOS Member Agency interfaces and security needs.

2.1.2 Purpose and Scope of Catalogue Interoperability

CEOS promotes the interoperability of space agency catalogues through the definition and development of interoperability concepts. By enhancing the standardization of EO data and information management services, CEOS enables the services to be more accessible and usable to data providers and data users worldwide. Catalogue interoperability also extends beyond just the members of CEOS in promoting data access within a wider community of EO data providers and eventually to non EO data providers.

2.1.3 Design Approach: CIP Space and ICS

The PTT design approach considers catalogue interoperability as the loose coupling of a federation of existing catalogue systems using a common protocol. The approach provides users the services and metadata available at all sites regardless of which site the user established a connection with. The

Catalogue Interoperability Protocol (CIP) standardizes the services needed for interaction between users and catalogues. The Interoperable Catalogue System (ICS) is a reference design which uses CIP as the common protocol between data providers and users. The objective of implementing CIP and ICS is to provide more users with access to more data more easily.

The CIP domain can be seen as a virtual 'CIP space' (see Figure 2-1) within which CIP messages, consisting of requests and responses, are exchanged between architectural elements. To support transparent access to multiple catalogues, a three tier structure was used to design the CIP space. Clients exchange messages with a middleware layer which in turn interacts with multiple catalogue servers. The middleware provides the routing and translation services to allow client requests to be presented at the multiple heterogeneous catalogues. The middleware is of two types of elements: Retrieval Managers and Translators. Retrieval Managers provide an access point for clients and route the requests to the various servers. Translators, bound with the clients and servers, translate CIP to and from the native protocol of the client or server. Future client and server developments may use CIP directly and hence not require translators. This is indicated by the shading in Figure 2-1.

This approach supports a diversity of clients, and servers. Clients may be used directly by a human user or may be an agency system acting on behalf of a user. Depending on the design of an existing catalogue system, services may be provided by different servers and translators. Because the routing service provided by a Retrieval Manager is independent of the type of service, separate translators may be provided for inventory, browse, ordering, and user profiles. This architecture is also applicable for small data providers, such as university research groups, who are unable to provide a Retrieval Manager at their site but still wish to join the CIP domain. Their local catalogue inventory can be made available to the CIP community by the inclusion within another agency's Retrieval Manager.

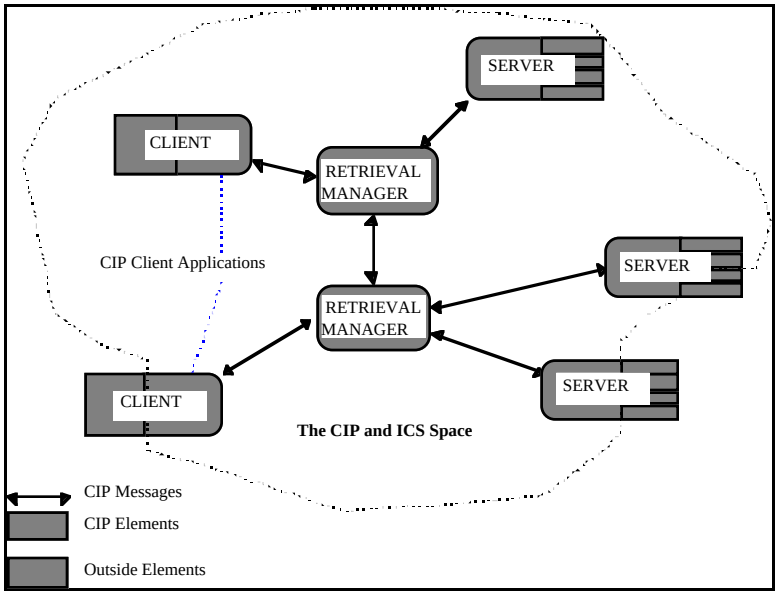

Figure 2-1 CIP Space

CIP space is a protocol centric view of catalogue interoperability and provides for the loosest coupling needed to achieve interoperability. A range of design solutions is permitted by the CIP space. To provide for a higher degree of uniform services at the cost of additional agreements between agencies, additional design for interoperability is defined in the ICS design document. The additional design definitions

pertain to the allocation of functionality and data amongst components, agreement on an underlying communication protocol, and agreement on how to conduct distributed system management of ICS. The difference between CIP Space and the ICS is depicted in Figure 2-2. CIP Space is defined by those CEOS agencies and other federations and organizations which provide catalogue services using CIP. Within CIP space are those CEOS agencies which provides services, communications and systems management compatible with the ICS design. Note that other federations may choose to use the ICS design as the basis for their federation.

descriptors for data products, guide documents, or other collections. In addition to the value of collections for presentation of data organization to users, collections provide the mechanism for routing distributed searches. When a collection contains both local and remote members, the Retrieval Manager may search both the local collection as well as sending the search on to the remote site.

The collections concept permits a wide degree of flexibility in EO product organisation. A collection has members consisting of product, guide or other collections. A collection groups product descriptors either due to their physical location (i.e. database storage) or along the lines of subject themes that are deemed useful access points. The collection concept is visualized in Figure 2-3 below. The collections in the diagram are numbered so that their relationship can be easily seen; they do not represent the naming of collections in an actual implementation. The terminal collections (labeled '1.x') group the product descriptors (inventory entries) as is appropriate. As can be seen the collections can overlap each other and product descriptors can appear in more than one collection. Above the terminal level collections, there are non-terminal collections that group together any number of other collections. The grouped collections do not all have to be at the same hierarchical level and this grouping of collections can continue to any hierarchical level, with existing collections being included at any other arbitrary level. A non-terminal collection could group together terminal collections and other non-terminal collections (as the link between collections 3.1 and 1.5 shows). Also, a terminal collection could exist without a relationship to a higher collection (i.e. collection 1.9), or a non-terminal collection could exist with no relationship to lower collections, in other words a collection without members (i.e. collection 2.5). Collection 1.9 can not be reached by a hierarchical search, but could be located if its URL was made public (an example of such a collection may be a hot collection or a collection under construction).

Collections can be used to group data together which have a similar semantic theme. Three types of collections are defined: Provider archive collection, Provider theme collection, User theme collection.

Provider archive collections: : This type of collection is likely to be created by data providers to organize their archives and facilitate access to the product descriptors (i.e. analogous to an inventory containing inventory entries). A provider archive collection could be a dynamic collection where new item descriptors are automatically added to the collection by the catalogue system as new archive entries are created, i.e. for data from an existing, operational EO satellite. Alternatively, a provider archive collection could be a static fixed collection of historical data that is no longer being added to as the source satellite is no longer operational. A provider archive collection includes only item descriptors that are local.

Provider theme collections: This type of collection may be set up by data providers who want to organize some of their data into groupings which differ from their provider archive collections (i.e. from the baseline inventory), for the convenience of their users, for example, based on the geographical area covered, the scientific discipline supported by the data, the instrument type, etc. The difference between provider theme and provider archive collections is that archive collections only contain homogeneous item descriptors, whilst theme collections may have item descriptors where each of the item descriptors has different attributes (e.g. local attributes). It is likely that there may be a large number that are homogeneous but not the complete collection. Provider theme collections can be terminal or non terminal dependent on how the data provider decides to organize their data or documents.

User theme collections: The third type of collection is likely to be created by an end user of EO products that has created and populated the collection to obtain a single source of thematic information. This will then enable further analysis or easy access by themselves or other users. These are collections of potentially quite disparate item descriptors of interest to a relatively small user community

researching a particular theme, i.e. in the example, the mid-west flood of 1993. There may be item descriptors from different image archives, in situ measurement archives and bibliographic archives of relevant papers. The members of a thematic collection will in general be formed from the results of a series of filtered searches to build up a set of descriptors.

Note that these category definitions are not mandatory for the CIP to operate, but help to distinguish collection categories for discussion purposes. The CIP does not distinguish between the categories (the Retrieval Manager does, however) and the same CIP search and retrieval services are applicable to all three. The Retrieval Manager does make use of collections for routing of distributed queries. Standardization of collection definitions is provided as part of the ICS design.

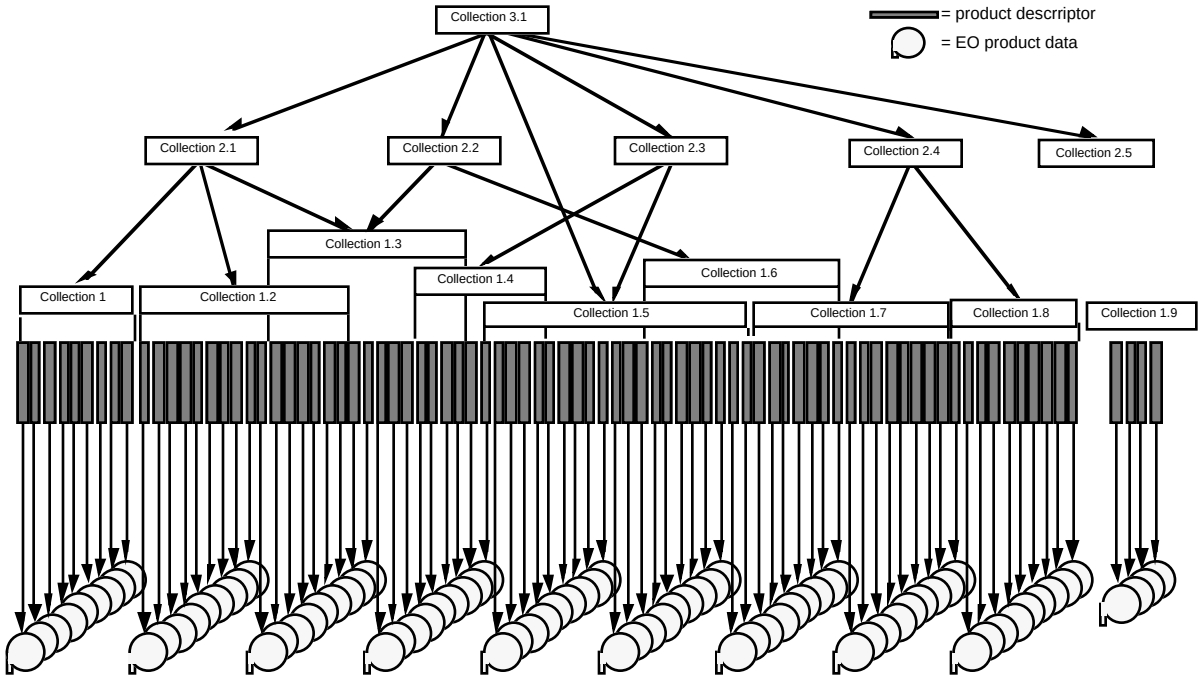

Figure 2-3 The concept of "collection"

2.1.5 CIP as a Z39.50 Profile

Based on a set of user requirements and an analysis of existing communication standards, Z39.50 was selected as the base protocol for CIP. CIP has exploited and extended the services of Z39.50 to provide distributed searching, extensions to attribute set definitions, and the definition of a secure ordering service. The Z39.50 protocol is designed for information search and retrieval within a generic domain which, together with the powerful services and data structures it supports, makes it an ideal basis of an EO domain search and retrieval protocol. CIP is a profile of Z39.50, i.e. it defines the use of the Z39.50 facilities within the CIP domain and defines the attributes that are used to search and present EO information. Other Z39.50 profiles include GILS, GEO and the Digital Collections Profile. CIP extends Z39.50 for distributed searching by supporting the collection data model discussed in section 2.1.4 which allows hierarchies of related collections to be constructed and searched.. Additional support for

compatibility is provided by the requirement that Retrieval Managers should be able to support access by any Z39.50 Version 3 compatible client.

CIP defines three types of searches which a CIP user may request:

- Collection Search: finds collections of interest without searching collections containing products
- Product Search: finds individual product descriptors which may eventually lead to the order of an actual product.
- Guide Search: finds guide descriptors which provide additional information (i.e. documents) describing collections or products.

Additionally, the user may request that the search be contained locally to the target Retrieval Manager (i.e., a local search), or request that the search be propagated to other Retrieval Managers based on the collections (i.e., a distributed search).

2.1.6 Browse Data and CIP

Browse data helps users to evaluate EO products. Browse data are typically reduced resolution or summary data versions derived from the EO product data itself. Browse data are delivered to the user via two different mechanisms, dynamically over the network during a user query session, and as an EO product order. The second case allows users to order the Browse data from an archive system to be delivered separately from their query session. This means that the user can then store and access the data locally rather than dynamically over a network. It is important to note that although most catalogue systems will provide some form of reduced data retrieval, it is not a mandatory CIP service. The form and content of browse data is dependent on the nature of the associated EO data and the data selection criteria necessary for a science discipline to evaluate the EO data. Browse data in the CIP is seen as one of the following forms:

- Browse attribute - simple attribute containing the actual Browse data.
- Browse compound - compound attribute containing attributes describing the Browse data, including the simple Browse data object.
- Browse product descriptor/product - handled as in product descriptors/products.

2.1.7 Product Ordering and Security

CIP includes an ordering method including the ability to specify order options and includes provisions for authentication and non-repudiation of orders. A user can retrieve the order options associated with a product, where order options may be processing as well as packaging options. CIP allows the local order handling system to define the order options without attempting to define an all encompassing order options standard. The user can request a quote for a specific order and submit the order. The order process is monitored by the Retrieval Manager and can be queried later by the user to determine the status of the order. To support ordering of data for which a user must have privileges or for orders which the user will be charged, a authentication scheme has been defined. The authentication supports digital signatures using either a shared (symmetric) key approach or an public (asymmetric) key approach. Authentication allows the Retrieval Manager to identify the user with an appropriate level of confidence and enables the Retrieval Manager to log the authenticated user requests to provide non-repudiation. The CIP security approach avoids the need to transfer password information over the network. Future enhancements to CIP anticipate the ability to support the transfer of financial information to support billing.

2.1.8 Implementation of CIP and ICS

The first implementations of ICS Elements have already been completed or are underway through ESA, EC/CEO, DLR, and NASA. The development of ICS will be accomplished in conjunction with other CEOS teams in particular on the topics of client and WWW functionality, browse, system management and network issues. Operational systems which are in preparation and which intend to implement CIP are underway by the following organisations: EC/CEO and ESA. Other organizations considering the use of CIP are NASA and the CCRS in Canada. Lessons learned from these implementations will be used to refine CIP and the ICS. A broad installation base of CIP systems can be expected, not because all catalogues have necessarily to be globally interoperable, but because CIP contains a consolidation of concepts critical to catalogues even in a more closed implementation.

2.2 Architecture Views

This System Design Document is developed based on a system architecture approach to system design [R6]. A system of the size and complexity of ICS will have multiple users, developers, operators, and

This page intentionally left blank.